

# A Model of Manual Control with Perspective Scene Viewing

Barbara T. Sweet\*

*NASA Ames Research Center, Moffett Field, CA, 94087, USA*

A model of manual control during perspective scene viewing is presented, which combines the Crossover Model with a simplified model of perspective-scene viewing and visual-cue selection. The model is developed for a particular example task: an idealized constant-altitude task in which the operator controls longitudinal position in the presence of both longitudinal and pitch disturbances. An experiment is performed to develop and validate the model. The model corresponds closely with the experimental measurements, and identified model parameters are highly consistent with the visual cues available in the perspective scene. The modeling results indicate that operators used one visual cue for position control, and another visual cue for velocity control (lead generation). Additionally, operators responded more quickly to rotation (pitch) than translation (longitudinal).

## Nomenclature

$D_X$	longitudinal position of scene feature in world coordinates, eyeheights
$D_Y$	lateral position of scene feature in world coordinates, eyeheights
$D_Z$	vertical position of scene feature in world coordinates, eyeheights
$F_x$	describing function model of operator control output to longitudinal position
$F_\theta$	describing function model of operator control output to pitch attitude
$\hat{F}_x$	describing function measurement of operator control to longitudinal position
$\hat{F}_\theta$	describing function measurement of operator control to pitch attitude
$H_C$	controlled element dynamics
$H_P$	human operator dynamic element
$H_x$	human operator dynamic element to longitudinal position
$H_\theta$	human operator dynamic element to pitch attitude
$I_h$	horizontal image coordinate
$I_v$	vertical image coordinate
$K_x$	gain parameter
$K_\gamma$	gain parameter for position
$K_\beta$	gain parameter for velocity
$L$	imaging device focal length
$P_{yy}$	power spectral density of $y(t)$
$P_{yz}$	cross power spectral density of $y(t)$ and $z(t)$
$r$	remnant
$s$	Laplace transform variable
$SE$	standard error
$t$	time, s
$W_{[A,B]}$	sensitivity parameter for visual cue A relative to state variable B
$X, x$	nonlinear, linearized operator longitudinal position
$Y, y$	nonlinear, linearized operator lateral position
$Z, z$	nonlinear, linearized operator vertical position
$V$	vertical displacement of a scene feature in the image plane

\*Aerospace Engineer, Human Systems Integration Research Division, MS 262-2, Senior Member

$\Delta V$	vertical displacement between two scene features in the image plane
$H$	horizontal displacement of a scene feature in the image plane
$\Delta H$	horizontal displacement between two scene features in the image plane
$S$	component of displacement along a line of splay in the image plane
$\alpha$	angle of a line of splay, rad
$\delta$	control input
$\chi^2$	chi-square function
$\Theta, \theta$	nonlinearized, linearized pitch attitude, rad
$\Phi, \phi$	nonlinearized, linearized roll attitude, rad
$\Psi, \psi$	nonlinearized, linearized heading attitude, rad
$\Gamma, \gamma$	nonlinear, linearized visual cue for position
$B, \beta$	nonlinear, linearized visual cue for velocity
$\tau$	time delay, s
$\tau_\theta$	differential time advance for pitch attitude $\theta$ , s
$\omega_N$	natural frequency of neuromuscular dynamics, rad/s
$\zeta_N$	damping of neuromuscular dynamics
$\omega_L$	lead equalization break frequency, rad/s
$\omega_C$	crossover frequency, rad/s
$\phi_M$	phase margin, deg

## I. Introduction

THIS paper describes a model of manual control in which the operator is using a perspective display; a perspective display is a two-dimensional depiction of a three-dimensional scene. Two of the main models of manual control are the Crossover Model<sup>1</sup> (CM) and the Optimal Control Model<sup>2</sup> (OCM). Manual control models have been primarily developed using compensatory and pursuit displays, as opposed to perspective displays.

A small number of researchers have extended the manual control methodologies by combining the OCM of the human operator with models of perspective-scene viewing. Grunwald and his colleagues have studied manual control extensively using perspective scenes for a variety of tasks and display types.<sup>3-12</sup> Zacharias developed general models of perspective-scene viewing using the OCM,<sup>13,14</sup> which have been applied to the analysis and design of simulator visual cues.<sup>15,16</sup> Wewerinke applied OCM techniques to examine the visual cues necessary for glideslope control in the landing task.<sup>17,18</sup> These OCM approaches have typically represented the perspective scene viewing with a linear combination of the states, where the combinatorial weights are governed by the perspective scene features. ‘Measurements’ obtained from the perspective scene are used as inputs to a state estimator, which provides a reconstruction of the system state to be used by the operator to control the vehicle.

The approach described in this paper is different than this previous work in two aspects. First, the CM, rather than the OCM, is used to represent the control actions of the human operator. Second, visual cues in the perspective scene are used as input to the model, rather than a reconstructed vehicle state. This is not to suggest that the human operator is not aware of the state of the vehicle – it is instead proposed that experienced operators learn to recognize the visual cues in the environment, or sight picture, that correspond to a particular desired vehicle condition. It is proposed that when such a nominal sight picture exists, that the operator uses the vehicle controls to maintain the nominal sight picture, rather than feeding back and controlling a full reconstruction of the vehicle state. This approach, and some of the research that inspired this approach, will be elaborated upon in the next section.

Relatively little previous work has been done in which the CM has been used to describe tasks in which a perspective display is used. Johnson and Phatak<sup>19</sup> combined the CM with an analysis of the characteristics of one visual element in a particular perspective display, and showed that the identified parameters were consistent with the perspective scene characteristics. Mulder examined the variations in identified CM parameters resulting from variations in design of a tunnel-in-the-sky display, but did not directly model the perspective scene characteristics.<sup>20</sup> In this paper, a modeling methodology is described which combines a model of perspective-scene visual processing with the CM. An experiment which was performed to develop and validate the model is described. The results of the experiment are used to infer the visual-cue usage strategy of the operators. The work described in this paper represents a significant expansion of the previous

body of knowledge; the previous work by Johnson and Phatak<sup>19</sup> modeled the use of one visual cue, for one operator, and one display condition. In this paper, the modeling methodology is developed in a way that can be generally extended to any task or display. Then, the model is applied to a particular task, and experimental validation is done with multiple operators and display types.

## II. The Visual Cue Control Model

In this section, the Visual Cue Control Model (VCCM) will be developed, first at a conceptual level, then in more detail for a particular task. The conceptual development consists of three parts: 1) discussion of other relevant research, 2) description of visual cues in the perspective scene, and 3) incorporation of visual cues into the CM.

### A. Previous Research

In the field of ecological psychophysics, extensive research has been conducted regarding how humans perceive and use information in a visual scene to accomplish self-motion (Warren and Wertheim edited a comprehensive compilation of research in this field<sup>21</sup>). The origins of ecological psychophysics date back to World War II, and research that James Gibson performed in an attempt to determine characteristics that could predict which pilot candidates would be successful.<sup>22</sup> After the end of the war, Gibson wrote about his experiences in a book that laid the foundations for the field of ecological psychophysics:<sup>23</sup>

Many tests were devised but none of them predicted a prospective flier's success or failure at this task. Many suggestions for training were made but none of them made the performance substantially easier. Toward the end of the war it began to be evident to psychologists working on problems of aviation that the usual approach to the problem of depth-perception was incorrect. Experiments needed to be performed *outdoors*. The stimuli judged ought to those of a *natural environment*.

In this spirit, many researchers have studied the visual cues that are useful to a pilot using visual references. One taxonomy that can be used to categorize visual cues is static versus dynamic. Static cues are those available in a discrete or momentary time step (Gibson used the term "momentary stimulation"). For example, the position of the horizon relative to the aircraft frame of reference of the pilot in the aircraft relate to the attitude of the aircraft. Linear perspective elements can inform the observer of depth and distance information; textures and spatial gradients of texture, as well as lighting and gradients of shading, specify distance and shape (and/or orientation) of surfaces in the world. Size of familiar objects is also a cue useful for determining depth. Discontinuities in both texture and linear elements can inform the observer of a discontinuity in depth, or an edge.

Dynamic cues can be utilized as well. Optic flow is defined as the pattern of visual movement in visual stimulation resulting from movement of the observer. Gibson considered the characteristics of optic flow that resulted from aircraft path and velocity.<sup>24,25</sup> One characteristic of the optic flow commonly used by pilots is locating the point of optical expansion to determine the landing location of the aircraft on the ground. In fact, Langschwiech discussed this cue in his 1944 flight training handbook.<sup>26</sup> A pilot on approach can ensure that he or she will land at the desired location by adjusting the aircraft controls to keep the point of optical expansion on the threshold of the runway; specific knowledge of glidepath (in degrees) is not necessary to perform the task.

This discussion has not been with the intent of providing an exhaustive or complete set of cues useful for visually-controlled flight; rather, it is intended to give examples of the many types of cues that can be identified and utilized. *A visual cue is, in essence, any definable feature or characteristic of the visual scene.* It can be the position or orientation of an element in the scene, a function of an area of the scene (such as a spatial gradient of texture elements). It can be a temporally specified features such the direction and/or magnitude of optic flow, or a discontinuity of optic flow, or the rate at which texture elements cross a boundary (such as the bottom edge of the scene).

Detailed descriptions of particular cues will be given in considering the example task below; at this point, the development will focus on an arbitrary visual cue definition, and the methodology used to derive the characteristics necessary for incorporation in the CM.

## B. Visual Cue Definition

While the previous section provided qualitative descriptions of visual cues, which are useful to understand the concept of visual cues, a quantitative description of the visual cue is necessary for modeling. This requires quantitative description of the perspective scene; specifically a mathematical description of the transformations which govern the image formation. Factors that can affect the perspective scene include 1) the locations of scene features, 2) the location and orientation of the imaging device (including vehicle state), and 3) the imaging device characteristics.

The scene features descriptions are typically available relative to a fixed, or world, coordinate system. Another coordinate system fixed to the vehicle being controlled is useful to describe the motion of the vehicle (in which the imaging device is located) relative to this fixed coordinate system. Eqs. 30 through 32 in Appendix A define the transformation from the location of a feature, imaging characteristics, and vehicle state, into image coordinates. The visual cue can be defined as a function of the image content. If, for example, we define a *nonlinear* visual cue  $\Lambda$  to be:

$$\Lambda = G_{image}(I_h, I_v) = G_{world}(L, X, Y, Z, \Theta, \Phi, \Psi, D_X, D_Y, D_Z) \quad (1)$$

$G_{image}(\bullet)$  represents an arbitrary function of the image-plane coordinates  $(I_h, I_v)$ .  $G_{world}(\bullet)$  represents the same function but expressed as a function of the focal length  $L$ , the position and orientation of the imaging device  $(X, Y, Z, \Theta, \Phi, \Psi)$ , and position of the scene feature  $D_X, D_Y, D_Z$ . It is obtained by substituting the expressions for  $I_h$  and  $I_v$  found in Eqs. 30 and 31 into the function  $G_{image}(I_h, I_v)$ . The parameters  $D_X, D_Y$  and  $D_Z$  and  $L$  are fixed for a particular feature and imaging geometry; the remaining variables are the vehicle states. Note that the arbitrary function  $G_{image}(\bullet)$  could be much more complex; it could be, for example, the summation of optical flow direction within certain boundaries in the image. The previous section provided many examples of possible visual cues.

The transformation between these vehicle states and the image-plane coordinates is nonlinear. A linear relationship is desired for incorporation with the quasi-linear CM, since this model describes the linear input/output relationships of the human operator. If, for example, it is assumed that the operator is trying to control  $X$  (longitudinal position), the *linearized* visual cue for that state,  $\lambda$ , can be defined as:

$$\lambda = \left. \frac{d\Lambda}{\partial\Lambda/\partial X} \right|_{X=X_0, Y=Y_0, Z=Z_0, \Theta=\Theta_0, \Phi=\Phi_0, \Psi=\Psi_0} \quad (2)$$

where

$$d\Lambda = \frac{\partial\Lambda}{\partial X}dX + \frac{\partial\Lambda}{\partial Y}dY + \frac{\partial\Lambda}{\partial Z}dZ + \frac{\partial\Lambda}{\partial\Theta}d\Theta + \frac{\partial\Lambda}{\partial\Phi}d\Phi + \frac{\partial\Lambda}{\partial\Psi}d\Psi \quad (3)$$

In this definition of the linearized cue  $\lambda$ , the differential of the nonlinear cue  $\Lambda$  is *normalized* with  $\partial\Lambda/\partial X$  to create one-to-one correspondence between the linearized cue  $\lambda$  and the longitudinal position; as a result,  $\lambda$  is expressed in units of longitudinal position. This was done to simplify incorporation into the CM.

By defining some additional terms, these equations can be simplified to provide a visual cue that is a linear combination of the vehicle states. We define the Sensitivity Parameter,  $W_{[\Lambda, Y]}$ , to be:

$$W_{[\Lambda, Y]} = \left. \frac{\partial\Lambda/\partial Y}{\partial\Lambda/\partial X} \right|_{X=X_0, Y=Y_0, Z=Z_0, \Theta=\Theta_0, \Phi=\Phi_0, \Psi=\Psi_0} \quad (4)$$

Other sensitivity parameters are defined similarly:  $W_{[\Lambda, \Theta]} = (\partial\Lambda/\partial\Theta)/(\partial\Lambda/\partial X)|_{X=X_0, \dots}$ , etc. We also define the linearized states  $x = dX$ ,  $y = dY$ , etc. With these definitions, we can rewrite Eq. 3:

$$\lambda = x + W_{[\Lambda, Y]}y + W_{[\Lambda, Z]}z + W_{[\Lambda, \Theta]}\theta + W_{[\Lambda, \Phi]}\phi + W_{[\Lambda, \Psi]}\psi \quad (5)$$

Given any definition of a nonlinear visual cue  $\Lambda$ , the expression for a linear visual cue  $\lambda$  can be derived that is simply a weighted sum of the vehicle states. The weightings are the sensitivity parameters as defined in Eq. 4, and are a function of the visual cue definition.

### C. Incorporation of Visual Cues into the Crossover Model

Much of the development of the CM was done using compensatory displays, in which the error between the desired state and the actual state is displayed to the operator. The perspective-scene viewing situation is fundamentally different from the compensatory-display viewing situation in two ways:

- 1) the perspective scene is affected by both controlled and uncontrolled states.
- 2) the process by which the vehicle states and feature coordinates are transformed into image coordinates is a non-linear transformation.

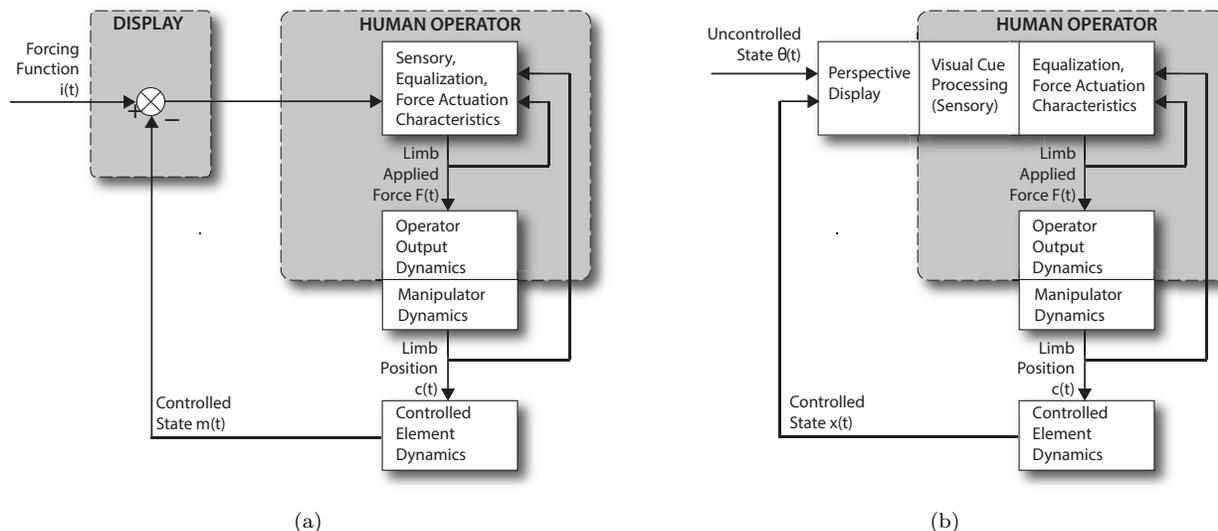


Figure 1. Single-loop manual control tasks with Compensatory (a) and Perspective (b) displays.

These two control-display viewing situations are shown conceptually in Fig. 1. Fig. 1a depicts a single-loop manual control system with a compensatory display.<sup>1</sup> In the compensatory display, only the error is presented to the operator. Fig. 1b depicts this single-loop control task accounting for perspective-scene viewing. In this case, the operator is performing regulation (or position-keeping) in the presence of disturbances, as opposed to tracking a commanded input.

While the systems depicted in Fig. 1 represent simplifications of the physical elements, they do not provide a good framework for isolation of the operator characteristics. To describe the input/output characteristics of the human operator depicted in Fig. 1a, McRuer and his colleagues developed the CM, a quasi-linear model of the human operator, consisting of a describing function plus remnant to represent the input-output characteristics of the operator (Fig. 2a). The describing function consisted of two elements: 1) a generalized describing function form, and 2) a series of adjustment rules for the describing function parameters. McRuer and his colleagues found that the adjustment rules were generally functions of the forcing function, controlled element dynamics, and frequency, and to a lesser extent a function of time and manipulator characteristics.<sup>1</sup>

As previously discussed, the inputs to the model are visual cues. This visual cue feedback, combined with the mathematical representation for visual cues derived in the previous section, lead to the block diagram representation of the VCCM shown in Fig.2b. While this block diagram shows only one controlled ( $x$ ) and one uncontrolled ( $\theta$ ) state, in general there will be a describing function for each state. The describing function relative to the controlled state(s) ( $H_x$  in Fig.2b) will be similar to the CM describing function ( $H_P$  in Fig.2a); this is the outcome of the previously stated assumption. The describing function relative to the uncontrolled state(s) ( $H_\theta$  in Fig.2b) will be a function of the visual cue(s) chosen to accomplish the control task.

Next the VCCM will be developed for a particular task that will later be used for experimental validation. The example derivation is presented at a level-of-detail that should be sufficient to allow the reader to apply the methodology to novel situations.

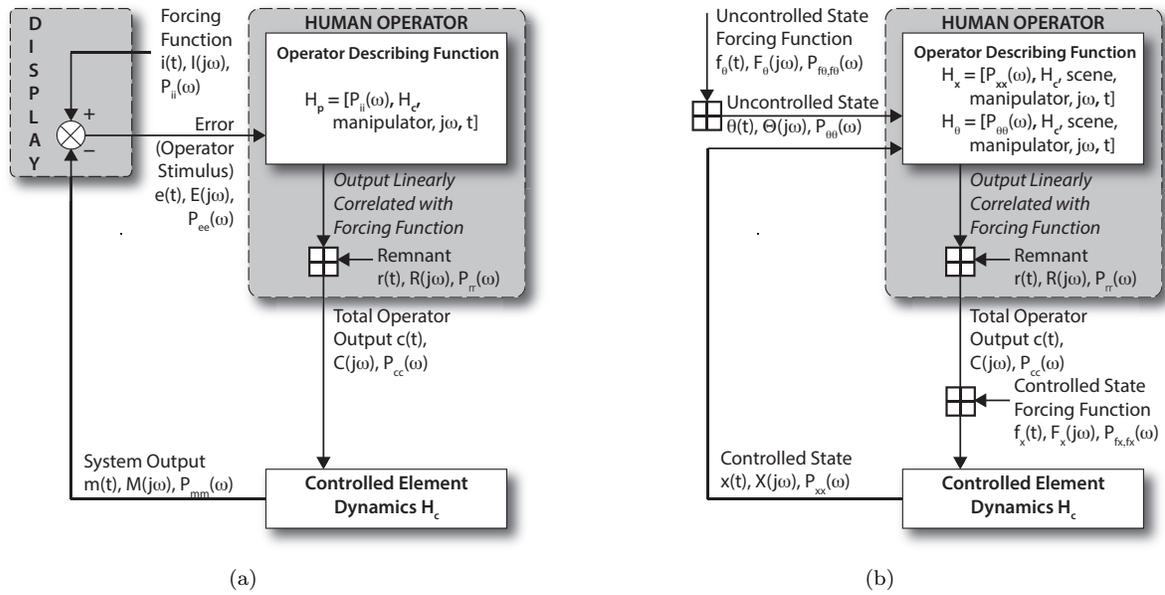


Figure 2. Equivalent block diagrams of the human operator in a manual control task, using Compensatory (a) and Perspective (b) displays.

## D. Example

### 1. Task

The task considered here is an idealized hover of a vehicle in the presence of disturbances. The only degrees of freedom allowed were longitudinal motion and pitch. The transfer functions representing the vehicle dynamics are taken to be:

$$x(s) = \frac{1}{s(s+0.2)}[\delta(s) + f_x(s)] \quad (6)$$

$$\theta = \frac{1}{s}f_{\theta}(s) \quad (7)$$

where  $\delta$  is the joystick displacement,  $x$  is the longitudinal position in units of eyeheights, and  $\theta$  is the pitch attitude in radians.  $f_x$  is a disturbance to the longitudinal acceleration in units of eyeheights/s<sup>2</sup>, and  $f_{\theta}$  is a disturbance in pitch rate in units of rad/s. Note that for this constant-altitude task, all distances are expressed relative to the altitude, or height, of the eyepoint of the operator. The lack of coupling between longitudinal position and pitch (as would occur in a helicopter) was a deliberate choice to simplify the modeling task.

### 2. Visual Cue Definition and Sensitivity Parameter Derivation

It is assumed that the operator finds visual cues that correlate with the desired state, in this case  $x$ . Given a nonlinear visual cue definition  $\Lambda$ , the linearized cue  $\lambda$  can be expressed as (Eq. 5):

$$\lambda = x + W_{[\Lambda, \Theta]}\theta \quad (8)$$

where  $W_{[\Lambda, \Theta]} = (\partial\Lambda/\partial\Theta)/(\partial\Lambda/\partial X)|_{X=X_0, \Theta=\Theta_0}$ . Note that the visual cue,  $\lambda$ , is simply a linear combination of  $x$  and  $\theta$ ; the weighting between the states,  $W_{[\Lambda, \Theta]}$ , specifies the relative contribution of  $\theta$  in the cue.

### 3. Describing Function Model Form

For the task described, the model form of the describing function will be developed assuming visual cues as inputs. However, it is necessary to first review the compensation the operator would adopt with a compensatory display, since the VCCM is based upon the assumption that the operator will adopt very similar compensation with a perspective display.

For the controlled-element dynamics in this example task ( $H_C = 1/[s(s + 0.2)]$ ), the CM would predict (in the region of the crossover frequency  $\omega_c$ ) that the product of the operator compensation  $H_P(s)$  and controlled element dynamics  $H_C(s)$  would be:

$$H_P(s)H_C(s) \approx \frac{\omega_c e^{-s\tau}}{s} \quad (9)$$

Accounting for the fact that the human can probably not generate 5 seconds of lead compensation<sup>27</sup> (necessary to cancel the pole at  $s = -0.2$  in the vehicle dynamics), one would expect the operator dynamics to take the approximate form:

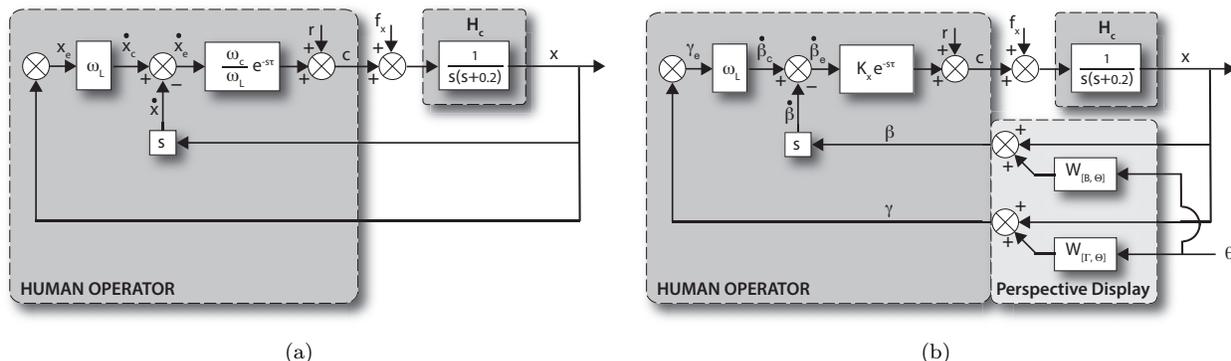
$$H_P(s) = \frac{\omega_c}{\omega_L} e^{-s\tau} (s + \omega_L) \quad (10)$$

where  $\omega_L$  should occur at a frequency below crossover, and at or above 0.2 rad/s ( $0.2 < \omega_L < \omega_c$ ).

Figure 3a shows a schematic diagram of this assumed compensation strategy. The transfer function between the control output and the controlled state would be:

$$\delta(s) = -H_P(s)x(s) + r(s) \quad (11)$$

The term  $r(s)$  is included in this transfer function and in the diagram; this represents remnant “injected” by the human operator into the control activity. Specifically, it is the control activity that is not linearly correlated with the input.



**Figure 3. Block diagrams of human operator performing the example task, using compensatory (a) and perspective (b) displays.**

Next the case of the operator using visual cues from a perspective display will be considered. For the controlled element dynamics considered in this example, the operator is required to generate a significant amount of lead compensation; this can be accomplished by feeding back both position and velocity. It is assumed that the operator can, and possibly will, use different visual cues for position and velocity. The operator could do this by using central, or foveal vision, to determine position of a cue, and parafoveal/peripheral vision to detect visual motion. We define  $\gamma$  to be the linearized visual cue for position, and  $\beta$  to be the linearized visual cue for motion, as follows:

$$\gamma(s) = x(s) + W_{[\Gamma, \Theta]} \theta(s) \quad (12)$$

$$\beta(s) = \dot{x}(s) + W_{[B, \Theta]} \theta(s) \quad (13)$$

Note that as defined, the visual cues  $\gamma$  and  $\beta$  are expressed in units of eyeheights, and the parameters defining the contribution of to the visual cue,  $W_{[\Gamma, \Theta]}$  and  $W_{[B, \Theta]}$ , are expressed in units of eyeheight/rad.

Now the usage of these cues will be examined with consideration of the expected operator compensation from the CM. Expressed in the time domain, Equations 10 and 11 become:

$$\delta(t) = \left( \frac{\omega_c}{\omega_L} \dot{x}(t - \tau) + \omega_L x(t - \tau) \right) + r(t) \quad (14)$$

Substituting  $\dot{\beta}$  for  $\dot{x}$ , and  $\gamma$  for  $x$ , and  $K_x$  for  $\omega_c/\omega_L$  we have:

$$\delta(s) = -H_x(s)x(s) - H_\theta(s)\theta(s) + r(s) \quad (15)$$

where

$$H_x(s) = K_x e^{-s\tau} (s + \omega_L) \quad (16)$$

$$H_\theta(s) = H_x(s) \left[ \frac{(W_{[B,\Theta]}s + W_{[\Gamma,\Theta]}\omega_L)}{(s + \omega_L)} \right] \quad (17)$$

The model resulting from inclusion of visual cues from the perspective display contains two separate describing functions. The first one,  $H_x$ , which is the operators' response to the longitudinal position  $x$ , is identical in form to the describing function that would be predicted for a compensatory display. The second describing function,  $H_\theta$ , is a function of not only  $H_x$ , but also the additional parameters  $W_{[\Gamma,\Theta]}$  and  $W_{[B,\Theta]}$ . These parameters are expected to be functions of the perspective display. Experimentally identified values of the parameters will later be compared with theoretically expected values for cues available in the perspective scene.

With the model form defined, the relationship between experimental measurements and describing functions can be derived. This is the subject of the next section.

#### 4. Describing Function Measurement

For experimental validation, it is necessary to relate the models to measurements. Time histories of the input/output variables  $x$ ,  $\theta$ ,  $\delta$ ,  $f_x$ , and  $f_\theta$  can be used to generate power-spectral and cross-spectral densities. Equations 6, 7, and 15 can be combined to produce the following relationships:

$$\frac{P_{\delta f_x}}{P_{x f_x}} = \frac{-H_x H_C P_{f_x f_x} - H_x H_C P_{\theta f_x} + P_{r f_x}}{H_C P_{f_x f_x} - H_x H_C H_\theta P_{\theta f_x} + H_C P_{r f_x}} \quad (18)$$

$$\frac{P_{\delta f_\theta}}{P_{\theta f_\theta}} = \frac{1}{1 + H_x H_C} \frac{-H_x H_C P_{u f_\theta} - H_x H_\theta P_{\theta f_\theta} + P_{r f_\theta}}{P_{\theta f_\theta}} \quad (19)$$

The input signals  $f_x$  and  $f_\theta$  can be made to have zero correlation with each other, and with ensemble averaging of measurements, the correlation between the remnant and the input signals can be minimized. With these conditions, the relationships become:

$$\frac{P_{\delta f_x}}{P_{x f_x}} = H_x \quad (20)$$

$$\frac{P_{\delta f_\theta}}{P_{\theta f_\theta}} = \frac{H_\theta}{(1 + H_x H_C)} \quad (21)$$

New terms will now be introduced to simplify later comparisons of the models and measurements:

$$\hat{F}_x = \frac{P_{\delta f_x}}{P_{x f_x}} \quad (22)$$

$$\hat{F}_\theta = \frac{P_{\delta f_\theta}}{P_{\theta f_\theta}} \quad (23)$$

$$F_x = H_x \quad (24)$$

$$F_\theta = \frac{H_\theta}{(1 + H_x H_C)} \quad (25)$$

The terms  $\hat{F}_x$  and  $\hat{F}_\theta$  refer to the experimental measurements. The terms  $F_x$  and  $F_\theta$  denote the parameterized models which will be developed to fit the measurements.

### III. Experimental Validation

A series of experiments was conducted to develop and validate the VCCM. The body of results from these experiments is too large to contain in this article; instead, a summary of the most significant results from one experiment are presented. Detailed descriptions of the experiments and results can be found in references 28 or 29.

## A. Protocol

A total of eight operators participated in the experiment. The display types used are shown in Fig. 4. The display types are called Grid (a), Perpendicular (b), Parallel (c), and Line (d). These display types were chosen to produce differences in the visual cues available to the operator. Each was rendered with a graphical field-of-view of 60 degrees (vertical) by 75 degrees (horizontal). Operators were told to control the longitudinal (fore-aft) position of the vehicle; they were also told that the only degrees of freedom were longitudinal and pitch displacement, and that other states would remain fixed.

The experimental apparatus was a part-task simulation hosted on an SGI Octane computer. A BG Systems JF3 joystick was used for the operator's control inputs. A 19-inch diagonal monitor was used, with a resolution of 1024 (vertical) by 1280 (horizontal) pixels. The display and joystick information were updated at a rate of 72 Hz. Each operator received extensive training runs before completing eight four-minute data runs with each display. Presentation order of the display conditions was counterbalanced between the operators. The time histories of the states  $x$ ,  $\theta$ , the control input  $\delta$ , and the disturbances  $f_x$ , and  $f_\theta$  were used to estimate the describing functions  $\hat{F}_x$  and  $\hat{F}_\theta$ , and the measurement standard errors  $SE(\hat{F}_x)$  and  $SE(\hat{F}_\theta)$ .<sup>30,31</sup>

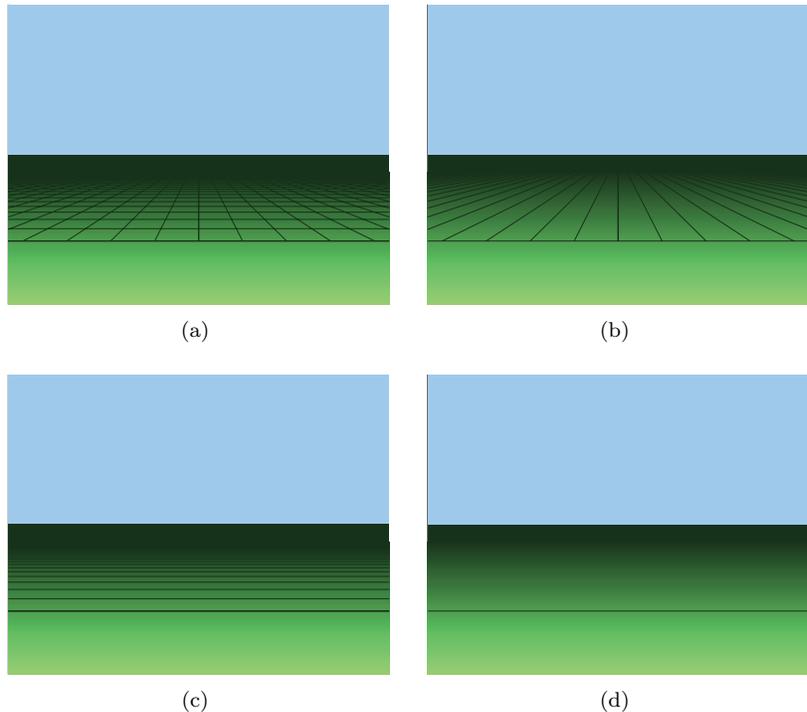
## B. Results

Describing function measurements for  $\hat{F}_x$  and  $\hat{F}_\theta$  were made using the methods previously described. Parameterized models were fit to the measurements using a maximum-likelihood estimate through minimization of the chi-square function,<sup>32</sup> defined as:

$$\chi^2 = \sum_{i=1}^{12} \left[ \frac{\text{real}[F_x(\omega_i) - \hat{F}_x(\omega_i)]^2}{SE[\text{real}(\hat{F}_x(\omega_i))]^2} + \frac{\text{imag}[F_x(\omega_i) - \hat{F}_x(\omega_i)]^2}{SE[\text{imag}(\hat{F}_x(\omega_i))]^2} \right] \quad (26)$$

$$+ \sum_{i=1}^{12} \left[ \frac{\text{real}[F_\theta(\omega_i) - \hat{F}_\theta(\omega_i)]^2}{SE[\text{real}(\hat{F}_\theta(\omega_i))]^2} + \frac{\text{imag}[F_\theta(\omega_i) - \hat{F}_\theta(\omega_i)]^2}{SE[\text{imag}(\hat{F}_\theta(\omega_i))]^2} \right] \quad (27)$$

The previously derived models in Eqs. 16 and 17 were based on the original CM, which was developed to describe manual control behavior in a limited frequency range. McRuer and his colleagues found that when the measurements in the region outside of crossover were sufficiently accurate, that a more complex model, which they named the Precision Model, was necessary to describe the data in the regions above and below crossover.<sup>33</sup> A similar result was found in fitting the data from these experiments; matching the high-frequency measurements required the addition of two terms. One term was a second-order neuromuscular dynamics term; the other was a different, shorter time delay in processing the rotational motion than in processing the translational motion. The resulting models were:



**Figure 4. The four display configurations tested. Grid (a), Perpendicular (b), Parallel (c), and Line (d).**

$$H_x(s) = \frac{K_x e^{-s\tau}(s + \omega_L)}{s^2 \omega_N^2 + 2s\zeta_N \omega_N + 1} \quad (28)$$

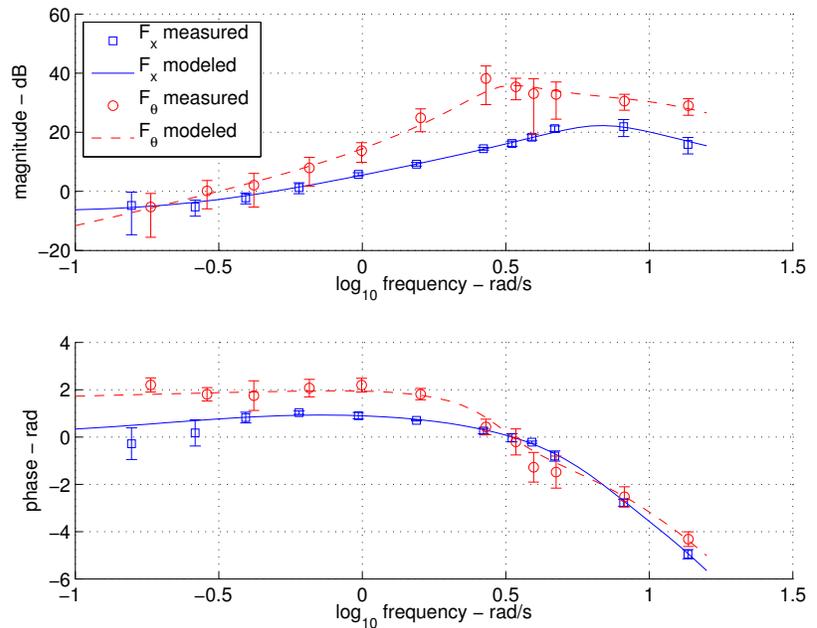
$$H_\theta(s) = H_x(s) e^{s\tau_\theta} \frac{K_\beta s + K_\gamma}{s + \omega_L} \quad (29)$$

Two simpler versions of this model were also considered; a seven-parameter model was tested for which the pitch motion time advance  $\tau_\theta$  was zero, and a six parameter model was tested in which the same visual cue was used for both position and velocity feedback (achieved by constraining  $K_\beta = K_\gamma$ ). Each parameter addition resulted in significant decreases in  $\chi^2$ . A rule of thumb is that when the chi-square value is approximately equal to the number of degrees of freedom (the number of measurements minus the number of parameters), the model is a moderately good fit to the data and further refinements are not warranted.<sup>32</sup> For this case, with 48 measurements for each operator and condition (two complex describing function measurements at twelve frequencies), a limiting value for model refinement would be 40.0. The fact that the average for the eight-parameter fit was nearly twice this amount would suggest that additional parameters could be warranted, but candidate ninth parameters did not result in better fits. The lack of further improvement is likely because the model was linear, and the perspective projection of the world coordinates into image coordinates is inherently nonlinear.

When measured in terms of magnitude and phase, the errors between the model and the measurements are modest. For the operator describing function to the longitudinal position,  $\hat{F}_x$ , the error between the model and measurement had a mean absolute value of 0.073 dB with a standard deviation of 1.41 dB, and a mean phase of -0.021 degrees with a standard deviation of 18.9 degrees. For the operator describing function to the longitudinal position,  $\hat{F}_\theta$ , the error between the model and measurement had a mean absolute value of -0.27 dB with a standard deviation of 1.73 dB, and a mean phase of -0.51 degrees with a standard deviation of 29.0 degrees. An example of the describing function data and resulting model fit for one operator and condition are shown in Fig. 5. Note that the standard errors are much smaller for the measurement of  $\hat{F}_x$  than  $\hat{F}_\theta$ ; this is a natural result of the fact that the operators are controlling position, not attitude.

Table 1 contains the mean and standard error of the seven model parameters, as well as crossover frequency ( $\omega_C$ ), phase margin ( $\phi_M$ ), and  $\chi^2$ , as a function of display type. A one-way ANalysis Of VAriance (ANOVA) was conducted on the model parameters, crossover frequency, and phase margin, to determine if the changes as a function of display type were statistically significant. The last two columns in Table 1 contain the F value resulting from the ANOVA, and the estimated probability ( $p$ ) that the observed effect of display is occurring from chance. Using a cutoff of  $p \leq 0.05$  for determining statistical significance,  $K_x$ ,  $\omega_N$ ,  $\zeta_N$ ,  $K_\gamma$ ,  $K_\beta$ ,  $\tau_\theta$ ,  $\omega_C$ , and  $\phi_M$  exhibited significant main effects from display type. Only the differences in time delay  $\tau$  and the lead break frequency  $\omega_L$  were not statistically significant.

The variables  $K_\gamma$ ,  $K_\beta$ ,  $\tau_\theta$ ,  $\omega_N$ ,  $\zeta_N$ ,  $K_\gamma$ ,  $K_\beta$ ,  $\tau_\theta$ ,  $\omega_C$ , and  $\phi_M$  are shown in Figure 6. The variable  $K_x$  is not shown independently because the primary effect of it is observed in the crossover frequency  $\omega_C$ . Paired t



**Figure 5. Example plot showing comparison of model fit with measurements. The operator was using the Parallel display for this measurement; standard error bars are shown.  $\chi^2$  for this model fit was 50.8.**

**Table 1. Model parameters. Values of  $p$  denoting statistical significance ( $p \leq 0.05$ ) are denoted with \*\*.**

parameter	Display					p
	Grid	Parallel	Perpendicular	Line	F	
$K_x$ (s • eyeheight)	$96.8 \pm 66.9$	$96.4 \pm 52.5$	$73.1 \pm 42.9$	$72.1 \pm 39.2$	5.75	0.0050**
$\omega_L$ (rad/s)	$0.569 \pm 0.268$	$0.565 \pm 0.273$	$0.591 \pm 0.279$	$0.505 \pm 0.227$	1.29	0.3050
$\tau$ (s)	$0.250 \pm 0.025$	$0.249 \pm 0.024$	$0.245 \pm 0.025$	$0.249 \pm 0.027$	0.43	0.7320
$\omega_N$ (rad/s)	$6.71 \pm 1.87$	$6.78 \pm 1.49$	$6.31 \pm 1.41$	$6.31 \pm 1.31$	4.11	0.0190**
$\zeta_N$ (rad/s)	$0.482 \pm 0.091$	$0.526 \pm 0.096$	$0.483 \pm 0.074$	$0.537 \pm 0.117$	4.54	0.0130**
$K_\gamma$ (s • eyeheight)	$8.32 \pm 3.96$	$10.07 \pm 2.69$	$10.12 \pm 2.70$	$12.68 \pm 2.60$	4.56	0.0130**
$K_\beta$ (s • eyeheight)	$3.07 \pm 1.69$	$4.38 \pm 1.57$	$6.34 \pm 1.21$	$8.07 \pm 0.43$	59.01	$< 0.0001$ **
$\tau_\theta$ (s)	$0.075 \pm 0.025$	$0.055 \pm 0.023$	$0.043 \pm 0.012$	$0.022 \pm 0.009$	13.60	$< 0.0001$ **
$\omega_C$ (rad/s)	$2.010 \pm 0.401$	$2.072 \pm 0.292$	$1.814 \pm 0.340$	$1.791 \pm 0.265$	8.38	0.0010**
$\phi_M$ (deg)	$33.45 \pm 8.10$	$30.68 \pm 6.55$	$35.76 \pm 6.15$	$36.46 \pm 5.68$	3.22	0.0440**
$\chi^2$	$93.96 \pm 45.50$	$74.48 \pm 24.43$	$89.73 \pm 33.48$	$75.13 \pm 43.06$	0.71	0.5580

tests were performed to determine whether differences observed between display conditions were statistically significant; the results are shown in Table 2.

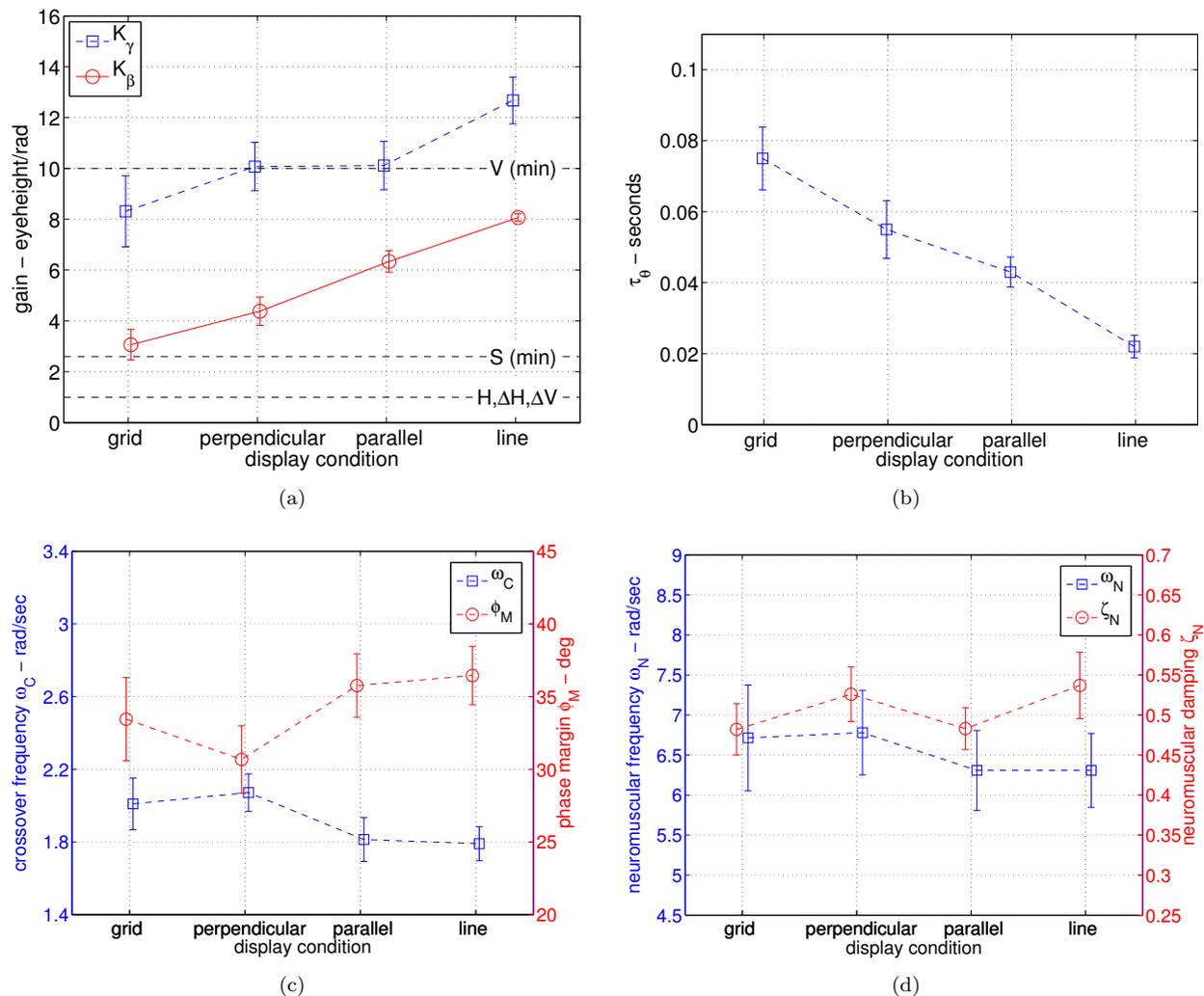
**Table 2. Paired t-tests statistical significance ( $p$  values) of the of the model parameters. The symbols #,  $\perp$ ,  $\parallel$ , and | are used to denote the grid, perpendicular, parallel, and line displays, respectively. As an example for the interpretation of the chart, the parameter  $\omega_C$  shows a significant difference between the grid and parallel display conditions, and the grid and line, but not the grid and perpendicular. Lack of statistical significance is shown with ‘-’.**

	$\omega_C$			$\phi_M$			$\omega_N$			$\zeta_N$		
	$\perp$	$\parallel$		$\perp$	$\parallel$		$\perp$	$\parallel$		$\perp$	$\parallel$	
#	-	0.013	0.023	-	-	-	-	-	-	-	-	0.035
$\perp$		0.017	0.001		-	0.013		0.003	0.015		-	-
$\parallel$			-			-			-			0.041
	$K_x$			$K_\gamma$			$K_\beta$			$\tau_\theta$		
	$\perp$	$\parallel$		$\perp$	$\parallel$		$\perp$	$\parallel$		$\perp$	$\parallel$	
#	-	0.042	-	-	-	0.035	0.001	$< 0.0001$	$< 0.0001$	-	0.005	0.001
$\perp$		0.003	0.022		-	-		0.001	$< 0.0001$		0.013	-
$\parallel$			-			-			0.003			0.001

## C. Discussion

### 1. Potential Visual Cues

For the example task, five potential types of visual cues, illustrated in Fig. 7, were considered that could provide the relevant state information to the operator. The derivation of the sensitivity parameters for each cue is contained in Appendix B. Symbols, analytical expressions and numerical values of the sensitivity parameters for each of the visual cues are shown in Table 3. One would anticipate that operators would choose visual cues that minimize the impact of the uncontrolled pitch attitude perturbations; this would imply the selection of a cue or cues that minimize the sensitivity parameter. Another factor that needs to be considered is whether the cue can be effectively perceived by the operator.



**Figure 6. Identified gain parameters  $K_\gamma$  and  $K_\beta$  (a), pitch time advance  $\tau_\theta$  (b), crossover frequency  $\omega_C$  and phase margin  $\phi_M$  (c), and neuromuscular frequency  $\omega_N$  and damping  $\zeta_N$ , averaged across subjects. Standard error bars are shown.**

**Table 3. Sensitivity parameters as a function of visual cue. All numerical values are expressed in units of rad/eyeheight. Display conditions in which the visual cue is present are denoted with an ‘x’.**

Visual Cue	Sensitivity Parameter			Display			
	Symbol	Expression	Value	Grid	Perpendicular	Parallel	line
$V$	$W_{[V,\theta]}$	$D_X^2 + 1$	$\geq 10.0$	x	x	x	x
$\Delta V$	$W_{[\Delta V,\theta]}$	1	1.0	x	x	x	x
$H$	$W_{[H,\theta]}$	1	1.0	x	x		
$\Delta H$	$W_{[\Delta H,\theta]}$	1	1.0	x	x		
$S$	$W_{[S,\theta]}$	$(D_Y^2 + D_X^2 + 1)/(D_Y^2 + 1)$	$\geq 2.6$	x	x		

First consider  $V$ , the vertical displacement of a single feature in the image. The sensitivity parameter of this cue is a function of the position of the scene feature, specifically  $W_{[V,\Theta]} = D_X^2 + 1$  eyeheight/rad. This parameter will be minimized by minimizing  $D_X$ , or using the feature closest to the observer. The minimum sensitivity parameter for this cue is  $W_{[V,\Theta]} = 10.0$  (when using the closest feature at  $D_X = 3.0$  eyeheights). This will be shown to be a relatively poor cue in relation to the other cues analyzed. However, this cue is easily perceived because 1) the line is clearly visible, and 2) the position of the line is being judged relative to a fixed image location (i.e. the screen edge). The visual cue  $V$  is available in all of the display configurations. The possible values of  $W_{[V,\Theta]}$  are shown as a function of image location in Figure 8a.

Next consider the characteristics of  $S$ , the component of the displacement of a feature parallel to the lines-of-splay. The expression for the sensitivity parameter is  $W_{[S,\Theta]} = (D_Y^2 + D_X^2 + 1)/(D_Y^2 + 1)$  eyeheight/rad. The value of  $W_{[S,\Theta]}$  will be minimized by minimizing  $D_X$ , and maximizing  $D_Y$ , which is achieved at the lower corners of the display, as shown in Figure 8b. For the features present in the displays, the lowest achievable value of  $W_{[S,\Theta]}$  is 2.6 eyeheight/rad. This is a superior cue to the absolute vertical displacement cue  $W_{[H,\Theta]}$  analyzed previously. It is anticipated that this cue would be used as a motion cue rather than a position cue by detecting *motion* in the direction of a line of splay. This cue is present in the Grid and Perpendicular displays, not the Parallel or Line displays.

The remaining three cues to consider,  $H$ ,  $\Delta H$ , and  $\Delta V$ , have a sensitivity parameter value of unity at every location in the display ( $W_{[H,\Theta]} = W_{[\Delta H,\Theta]} = W_{[\Delta V,\Theta]} = 1$  eyeheight/rad). To review,  $H$  is the horizontal displacement of a single feature in the image.  $\Delta H$  and  $\Delta V$  are, respectively, horizontal and vertical displacements between two scene features in the image. While a sensitivity parameter magnitude of unity is superior to the previously considered cues, these three cues may be of limited practical usage. First consider the two cues related to horizontal judgements, either the absolute horizontal displacement of a feature  $H$  or relative horizontal displacement between features  $\Delta H$ . Since the task is longitudinal position control, and the control effector moves in the longitudinal plane, there might be some practical difficulty in using a horizontal (i.e. lateral) displacement cue because of the lack of intuitive mapping between the cues; at best, one would expect control reversals to occur. Now consider the cue related to the relative vertical displacement between features ( $\Delta V$ ). Although this cue potentially has a better intuitive mapping to longitudinal position, one problem that might be encountered using this cue is the fact that the two features would both be moving because of the pitch motion. The constant motion of both features could make this cue difficult to perceive. The cue using relative horizontal displacements ( $\Delta H$ ) could be difficult to use for this reason as well.

In summary, three potential visual cues,  $H$ ,  $\Delta H$ , and  $\Delta V$ , have the best theoretical characteristics, but perceiving them accurately could be difficult. The worst cue to use would be  $V$ , although it is probably easily perceived.  $S$  would be a potentially good cue for motion perception, particularly if the best cues are difficult to perceive.

## 2. Visual Cue Identification

Now we will consider the pilot model parameters ( $K_\gamma$  and  $K_\beta$ ) that correspond to the two hypothetical visual cues ( $\beta$  for motion,  $\gamma$  for position) in the VCCM. Identified values of  $K_\gamma$  and  $K_\beta$  are shown in Table 1 and Figure 6a; the plot also shows the predicted values of the sensitivity parameters for the potential visual cues analyzed. The differences in these parameters were statistically significant, with probabilities of the differences occurring by chance of  $p = 0.013$  and  $p < .0001$  for  $K_\gamma$  and  $K_\beta$ , respectively. The identified parameters generally fall within the expected range for the available visual cues. In no case did an identified

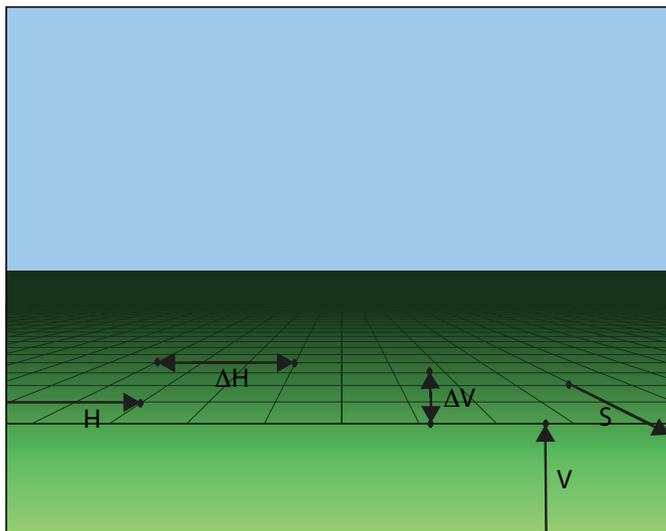
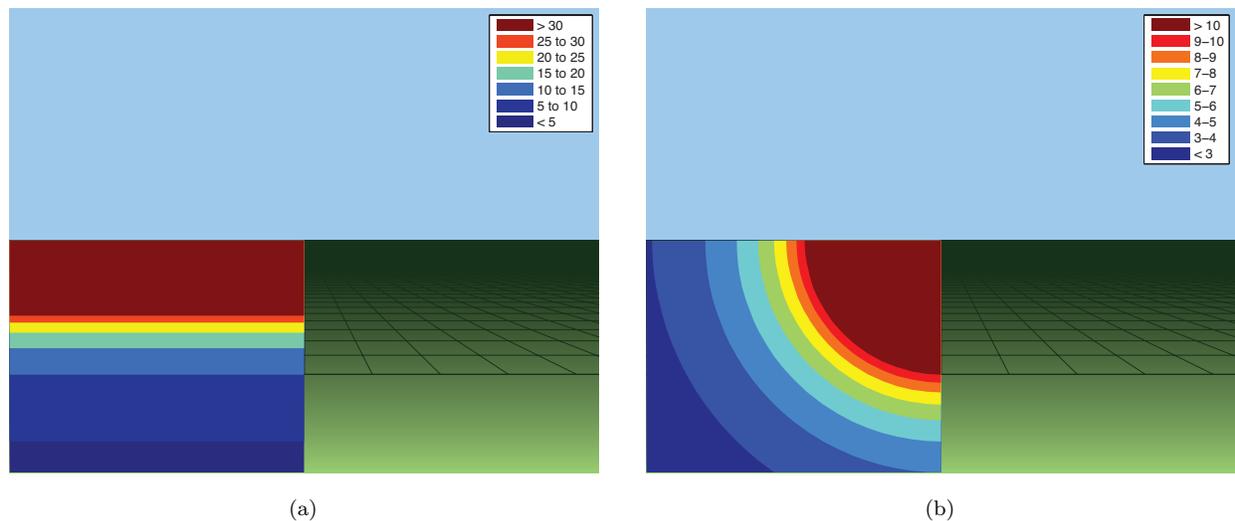


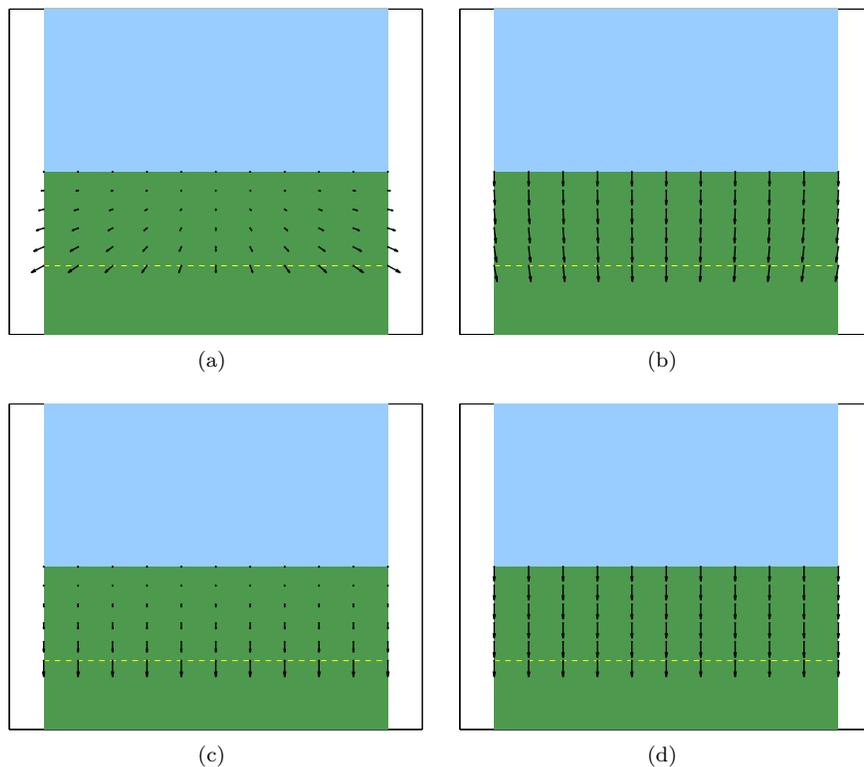
Figure 7. Examples of available visual cues in the displays tested.



**Figure 8.** Direction of image motion resulting from longitudinal displacement (a) and pitch displacement (b). Nominal position of a location 3.0 eyeheights in front of the observer is shown with a dotted line.

parameter go below unity, and most of the parameters are between 1 and 10 eyeheight/rad. Also note that the changes in display type produced systematic changes in the identified parameters.

Now let us examine these parameters in more detail, starting with  $K_\gamma$ , the identified gain on the visual cue  $\gamma$ . The identified  $K_\gamma$  was very close to the predicted value of 10 eyeheight/rad for  $W_{[V,\Theta]}$ . The difference in  $K_\gamma$  between the Grid and Line display conditions was statistically different, with lower values being obtained with the grid display. At the conclusion of the experiments, the operators were asked for verbal descriptions of their visual cue strategies. Many of them indicated that they were trying to keep the distance between the horizon and the line in the foreground constant (which would produce  $W_{[r,\Theta]} = 1$  eyeheight/rad), but that because of the motion of both lines they needed to also reference the foreground line (which would produce  $W_{[r,\Theta]} = 10$  eyeheight/rad). Values of  $K_\gamma$  less than 10 could be consistent with an attention-sharing strategy between the two cues, as indicated by the operators. It appears that the richer scene content in the Grid



**Figure 9.** Direction of image motion resulting longitudinal motion and pitch rate. Longitudinal motion is shown in (a) and (c) for the Grid/Perpendicular, and Parallel/Line displays, respectively. Pitch motion is shown in (b) and (d) for the Grid/Perpendicular, and Parallel/Line displays.

display provided operators with a greater ability to disregard the effects of pitch.

Now we will consider  $K_\beta$ , the identified gain on the visual cue  $\beta$  in the pilot model. The paired t-test comparisons revealed that the differences in this parameter between each display pair was statistically significant. Notably, the mean value of  $K_\beta$  when using the Grid display is close to 2.6 eye heights/rad, the value associated with the optimal use of  $S$  (motion along a line-of-splay). When interviewed, many operators verbally indicated that they were attending to their peripheral vision to “quicken” their responses. Interestingly, one operator achieved values better than what could be achieved using  $S$  with both the Grid and Parallel displays. When interviewed, this operator indicated that he had indeed discovered and used the “optimal” cue  $H$ ; he was controlling the horizontal displacement of a grid intersection (only available in the Grid and Parallel displays) relative to the image frame. This strategy would correspond to a sensitivity parameter  $W_{[H,\Theta]}$  of unity, very close to the identified value of  $K_\beta$  for this operator. The identified values of  $K_\beta$  were lower when using the displays containing the angle-of-splay cue  $S$ , the Grid and the Parallel display. The fact that operators achieved values lower than that expected for using visual cue  $V$  could suggest that operators could perceive the motion between lines ( $\Delta V$ ) to some extent; even in the Line display, operators can use the displacement between the horizon and the line. The fact that values of  $K_\beta$  were consistently lower than  $K_\gamma$  indicate that the visual motion processing capabilities of the operators yielded better differentiation between the effects of rotation and translation.

$\tau_\theta$  is another model parameter related to visual cue processing. This parameter is shown in Table 1 and Figure 6b, and was statistically significant ( $p < 0.0001$ ). This parameter describes a time advance in the use of pitch attitude  $\theta$  relative to the longitudinal position  $x$ . In all cases, pitch was acted upon slightly faster than longitudinal position. However, the amount of difference depended upon the presence of the lines-of-splay. Pitch time advance  $\tau_\theta$  with the Grid and Perpendicular displays was higher than with the Parallel and Line, as shown in the pairwise comparisons in Table 2. These differences are likely due to significant differences in how the longitudinal and pitch displacements affect the different display types. Figure 9 illustrates how longitudinal and pitch displacements affect the Grid and Perpendicular displays (a and b), and how these displacements affect the Parallel and Line displays (c and d). In the Grid/Perpendicular displays, longitudinal movement creates significant differences in displacement/optic flow magnitude and direction as a function of location in the image (Figure 9a). Pitch movement creates mostly uniform vertical displacement and optic flow (Figure 9b). In the Parallel/Line displays, the difference arising from the states is chiefly in the distribution of vertical motion (Figures 9c and 9d). It is likely that operators could disambiguate the effects of pitch and longitudinal motion more quickly with the Grid and Perpendicular displays, leading to increases in time advance.

Some of the parameters related to the primary control task (control of longitudinal position  $x$ ) were also statistically significant, although these were largely main effects of display condition rather than significant effects between display conditions. Crossover frequency  $\omega_C$  and phase angle  $\phi_M$  are shown in Figure 6c, and neuromuscular parameters  $\omega_N$  and  $\zeta_N$  are shown in Figure 6d. The crossover frequency was higher in the Grid and Perpendicular display conditions than with the Parallel and Line displays. It is likely that uncertainty in perceiving the controlled variable due to pitch variations caused operators to reduce their control gains.

## IV. Conclusions

The Visual Cue Control Model has been shown to be highly accurate at describing the human operator characteristics in performing a longitudinal control task with a perspective display. The model incorporates parameters that can be directly related to the use of visual cues in the perspective display. The experimental validation showed that 1) operators used two different cue sources for position and velocity feedback; 2) identified parameters for the position feedback visual cue  $K_\gamma$  and the velocity visual cue  $K_\beta$  corresponded closely with the hypothetical visual cues present in the different display conditions; and 3) pitch attitude was reflected in the operator output more quickly than longitudinal position, with this differential time increasing with display complexity. Two of the model parameters associated with visual cue usage,  $K_\beta$ , and  $\tau_\theta$ , demonstrated not only large main effects, but also significant differences as a function of display type in pairwise comparisons. Overall, the Grid and Perpendicular displays were associated with lower values of  $K_\beta$ , and higher values of  $\tau_\theta$ , than the Parallel and Line displays. The presence of the lines of splay appear to enable better ‘rejection’ of the pitch attitude variations, as well as faster processing of pitch attitude.

Although the ANOVA analysis demonstrated many significant main effects when averaged across opera-

tors, operators sometimes self-reported very different visual cue usage strategies. A follow-on (Monte Carlo) analysis is planned to identify the statistical properties of the parameters for individuals, in order to study individual differences.

The model shows great promise as a tool for determining optimum scene content, as well as development of training instructions for visually guided control tasks. This is evidenced by the fact that only one of the eight operators found and used one of the optimal visual cues. It should be possible to do a priori analysis of a task to determine the most effective cues, and use this as the basis for training. The model, when used to identify parameters in an experimental setting, was also useful in determining when a theoretically optimal cue was not being used effectively. This is an important feature for validation of perspective scene designs: it can be determined if the cues are being used as expected, and/or when a particular visual cue is ineffective.

The Visual Cue Control Model is generally extensible to any manual control task in which a perspective display is used. This technique incorporates the well-validated Crossover Model, and additional parameters associated with the perspective scene and visual cues. These additional parameters are readily identified in an empirical setting, and can be used to verify the visual-cue usage of the human operator.

## Appendices

### A. General Perspective Projection Transformation

The coordinate system  $(\hat{I}, \hat{J}, \hat{K})$  is inertial, assumed fixed in the world (see Fig. 10). The feature being imaged is located at  $\bar{D} = D_X \hat{I} + D_Y \hat{J} + D_Z \hat{K}$ , The operator eyepoint is located at  $\bar{P} = X \hat{I} + Y \hat{J} + Z \hat{K}$ . The orientation of the vehicle-fixed coordinate system  $(\hat{i}, \hat{j}, \hat{k})$ , relative to the inertial coordinate system, is described by the Eulerian angles<sup>34</sup>  $\Psi$  (heading),  $\Theta$  (pitch), and  $\Phi$  (roll). The vehicle coordinate system can be obtained from the inertial coordinate system through three sequential body-fixed rotations:  $\Psi$  about  $\hat{k}$ ,  $\Theta$  about  $\hat{j}$ , and  $\Phi$  about  $\hat{i}$ .

The transformation which relates the operator location and orientation  $(X, Y, Z, \Phi, \Theta, \Psi)$ , feature location  $(D_X, D_Y, D_Z)$ , and focal length  $(L)$ , to image coordinates,<sup>35</sup> is:

$$I_h = w \left[ (D_X - X)(-\sin \Psi \cos \Phi + \cos \Psi \sin \Theta \sin \Phi) \right. \\ \left. + (D_Y - Y)(\cos \Psi \cos \Phi + \sin \Psi \sin \Theta \sin \Phi) + (D_Z - Z) \cos \Theta \sin \Phi \right] \quad (30)$$

$$I_v = w \left[ (D_X - X)(\sin \Psi \sin \Phi + \cos \Psi \sin \Theta \cos \Phi) \right. \\ \left. + (D_Y - Y)(-\cos \Psi \sin \Phi + \sin \Psi \sin \Theta \cos \Phi) + (D_Z - Z) \cos \Theta \cos \Phi \right] \quad (31)$$

where  $w$  is defined as:

$$w = \frac{L}{(D_X - X) \cos \Phi \cos \Theta + (D_Y - Y) \sin \Phi \cos \Theta - (D_Z - Z) \sin \Theta} \quad (32)$$

### B. Visual Cue Sensitivity Parameter Derivation

For the example task (see Fig. 3), in which the only vehicle degrees-of-freedom are longitudinal position and pitch attitude, and vertical position is set to  $Z = 1.0$  eyeheight, the relationships defining the image-plane coordinates of a particular scene feature (Eqs. 30 through 32) simplify to:

$$I_h = \frac{LD_Y}{\left[ (D_X - X) \sin \Theta + \cos \Theta \right]} \quad (33)$$

$$I_v = \frac{L \left[ (D_X - X) \sin \Theta + \cos \Theta \right]}{\left[ (D_X - X) \cos \Theta - \sin \Theta \right]} \quad (34)$$

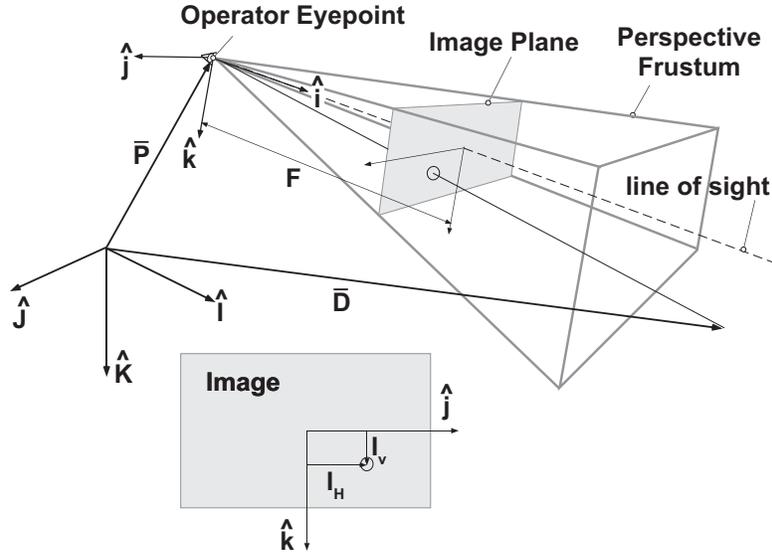


Figure 10. Geometry of the imaging process, relating vehicle state variables and scene feature locations in world coordinates to image coordinates.

in which  $L$  is the focal length,  $X$  is the longitudinal position and  $\Theta$  is the pitch attitude of the operator's vehicle reference frame, and  $D_X$  and  $D_Y$  are the longitudinal and lateral locations, respectively, of a scene feature (all expressed in units of eyeheights).

### 1. Vertical Location of a Feature in the Image

The visual cue  $V$  is defined as the vertical position, in the image, of a feature located at a longitudinal position  $D_X$  in the world:

$$V = I_v = \frac{L \left[ (D_X - X) \sin \Theta + \cos \Theta \right]}{(D_X - X) \cos \Theta - \sin \Theta} \quad (35)$$

The sensitivity parameter is defined as (from Eq. 4):

$$W_{[V,\Theta]} = \left. \frac{\partial V / \partial \Theta}{\partial V / \partial X} \right|_{X=X_0, \Theta=\Theta_0} \quad (36)$$

The partial derivatives are:

$$\frac{\partial V}{\partial X} = L \left[ \frac{-\sin \Theta}{(D_X - X) \cos \Theta - \sin \Theta} + \frac{\cos \Theta [(D_X - X) \sin \Theta + \cos \Theta]}{[(D_X - X) \cos \Theta - \sin \Theta]^2} \right] \quad (37)$$

$$\frac{\partial V}{\partial \Theta} = L \left[ 1 + \frac{[(D_X - X) \sin \Theta + \cos \Theta]^2}{[(D_X - X) \cos \Theta - \sin \Theta]^2} \right] \quad (38)$$

Evaluating at the linearization conditions of  $X_0 = 0$  and  $\Theta_0 = 0$  yields:

$$\left. \frac{\partial V}{\partial X} \right|_{X=0, \Theta=0} = \frac{L}{D_X^2} \quad (39)$$

$$\left. \frac{\partial V}{\partial \Theta} \right|_{X=0, \Theta=0} = L \left[ 1 + \frac{1}{D_X^2} \right] \quad (40)$$

The sensitivity parameter  $W_{[V, \Theta]}$  results from substituting equations 39 and 40 into 36:

$$W_{[V, \Theta]} = D_X^2 + 1 \quad (41)$$

## 2. Vertical Displacement between Two Features

The visual cue  $\Delta V$  is the vertical displacement between two features located at vertical positions  $I_{z1}$  and  $I_{z2}$  in the image, and longitudinal positions  $D_{x1}$  and  $D_{x2}$  in the world. Mathematically,  $\Delta V$  is expressed as:

$$\Delta V = I_{v2} - I_{v1} = L \left[ \frac{(D_{X2} - X) \sin \Theta + \cos \Theta}{(D_{X2} - X) \cos \Theta - \sin \Theta} - \frac{(D_{X1} - X) \sin \Theta + \cos \Theta}{(D_{X1} - X) \cos \Theta - \sin \Theta} \right] \quad (42)$$

Application of Equation 4 will result in a sensitivity parameter  $W_{[\Delta V, \Theta]} = 1$ .

## 3. Horizontal Displacement Between Two Features

The visual cue  $\Delta H$  is defined as the horizontal displacement between any two features in the image, located at horizontal positions  $I_h$  and  $I_v$  in the image, and positions  $(D_{X1}, D_{Y1})$  and  $(D_{X2}, D_{Y2})$  in the world.  $\Delta H$  is defined as:

$$\Delta H = I_{h2} - I_{h1} = L \left[ \frac{D_{Y2}}{(D_{X2} - X) \cos \Theta - \sin \Theta} - \frac{D_{Y1}}{(D_{X1} - X) \cos \Theta - \sin \Theta} \right] \quad (43)$$

The resulting sensitivity parameter is  $W_{[\Delta H, \Theta]} = 1$ .

## 4. Component of Motion along a Line-of-Splay

A line-of-splay is the line, in the image, formed by lines in the world that are parallel to the operator's direction of motion; several of these lines are indicated in Fig. 7. This cue consists of the component of displacement of a feature *along a line-of-splay*. This cue is considered because the motion of features, in the image, due to longitudinal motion of the vehicle, is along these lines-of-splay. The image motion resulting from pitch movement is largely vertical. Thus, if the operator attempts to control only motion that occurs along a line-of-splay, they would largely be controlling longitudinal motion. If the angle of the line-of-splay, relative to vertical, is defined to be  $\alpha$ , the component of that displacement parallel to the line-of-splay, defined as  $W$ , is:

$$W = I_h \sin \alpha + I_v \cos \alpha \quad (44)$$

$$W = L \left[ \frac{D_Y}{(D_X - X) \cos \Theta - \sin \Theta} \sin \alpha - \frac{(D_Y - Y) \sin \Theta + \cos \Theta}{(D_X - X) \cos \Theta - \sin \Theta} \cos \alpha \right] \quad (45)$$

The sine and cosine of  $\alpha$  can also be expressed in terms of world coordinates:

$$\sin \alpha = D_Y \cos \Theta / \sqrt{D_Y^2 \cos^2 \Theta + 1} \quad (46)$$

$$\cos \alpha = 1 / \sqrt{D_Y^2 \cos^2 \Theta + 1} \quad (47)$$

The sensitivity parameter for this cue can be shown to be:

$$W_{[S, \Theta]} = 1 + \frac{D_X^2}{(D_Y^2 + 1)} \quad (48)$$

## Acknowledgments

The author would like to thank Prof. Ron Hess at U.C. Davis, Prof. Dave Powell at Stanford, and Drs. Mary Kaiser and Leland Stone at NASA Ames for their encouragement, guidance and support in conducting this research. Thanks also go to Dr. Peter Zaal for reviewing this paper and providing valuable feedback, and to Dr. Martine Godfroy for performing statistical analysis (and associated interpretation) of the data.

## References

- <sup>1</sup>Mcruer, D. T., Krendel, E., and Reisener, W. J., "Human Pilot Dynamics in Compensatory Systems," Tech. Rep. AFFDL-TR-65-15, Air Force Flight Dynamics Laboratory Technical Report, 1965.
- <sup>2</sup>Kleinman, D. L., Baron, L. S., and Levison, W. H., "An Optimal Control Model of Human Response, Part I," *Automatica*, Vol. 6, No. 3, 1970, pp. 357–369.
- <sup>3</sup>Grunwald, A. J. and Merhav, S. J., "Vehicular Control by Visual Field Cues – Analytical Model and Experimental Validation," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 6, No. 12, 1976, pp. 835–845.
- <sup>4</sup>Grunwald, A. J. and Merhav, S. J., "Effectiveness of Basic Display Augmentation in Vehicular Control by Visual Field Cues," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 8, No. 9, 1978, pp. 679–690.
- <sup>5</sup>Merhav, S. J. and Grunwald, A. J., "Display Augmentation in Manual Control of Remotely Piloted Vehicles," *Journal of Aircraft*, Vol. 15, No. 3, 1978, pp. 182–189.
- <sup>6</sup>Grunwald, A. J., Robertson, J. B., and Hatfield, J. J., "Experimental Evaluation of a Perspective Tunnel Display for Three-Dimensional Helicopter Approaches," *Journal of Guidance and Control*, Vol. 4, No. 6, 1981, pp. 623–631.
- <sup>7</sup>Grunwald, A. J., "Tunnel Display for Four-Dimensional Fixed-Wing Aircraft Approaches," *Journal of Guidance, Control and Dynamics*, Vol. 7, No. 3, 1984, pp. 369–377.
- <sup>8</sup>Negrin, M., Grunwald, A. J., and Rosen, A., "Superimposed Perspective Visual Cues for Helicopter Hovering above a Moving Ship Deck," *Journal of Guidance*, Vol. 14, No. 3, 1991, pp. 652–660.
- <sup>9</sup>Grunwald, A. J. and Kohn, S., "Flight-Path Estimation in Passive Low-Altitude Flight by Visual Cues," *Journal of Guidance, Control and Dynamics*, Vol. 16, No. 2, 1993, pp. 363–370.
- <sup>10</sup>Grunwald, A. J. and Kohn, S., "Visual Field Information in Low-Altitude Visual Flight by Line-of-Sight Slaved Helmet-Mounted Displays," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 24, No. 1, 1994, pp. 120–134.
- <sup>11</sup>Grunwald, A. J., "Improved Tunnel Display for Curved Trajectory Following: Control Considerations," *Journal of Guidance, Control and Dynamics*, Vol. 19, No. 2, 1996, pp. 370–377.
- <sup>12</sup>Grunwald, A. J., "Improved Tunnel Display for Curved Trajectory Following: Experimental Evaluation," *Journal of Guidance, Control and Dynamics*, Vol. 19, No. 2, 1996, pp. 378–384.
- <sup>13</sup>Zacharias, G. L., "Visual Cueing Model for Terrain-Following Applications," *Journal of Guidance, Control and Dynamics*, Vol. 8, No. 2, 1985, pp. 201–207.
- <sup>14</sup>Zacharias, G. L., "An Estimation/Control Model of Egomotion," *Perception and Control of Self-Motion*, edited by R. Warren and A. H. Wertheim, Lawrence Erlbaum Associates, Hillsdale, 1990, pp. 425–459.
- <sup>15</sup>Baron, L. S., Lancraft, R., and Zacharias, G. L., "Pilot/Vehicle Model Analysis of Visual and Motion Cue Requirements in Flight Simulation," Tech. Rep. CR-3312, NASA, 1980.
- <sup>16</sup>Zacharias, G. L., "Modeling the Pilot's Use of Flight Simulator Visual Cues in a Terrain-Following Task," Tech. Rep. R8505, Charles River Analytics, 1985.
- <sup>17</sup>Wewerinke, P. H., "The Effect of Visual Information on the Manual Approach and Landing," NLR TR 80055, NLR, 1980.
- <sup>18</sup>Wewerinke, P. H., "Visual Scene Perception in Manual Control," *Journal of Cybernetics and Information Science*, Vol. 3, 1980, pp. 1–4.
- <sup>19</sup>Johnson, W. and Phatak, A., "Optical Variables and Control Strategy used in a Visual Hover Task," *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 1989, pp. 719–724.
- <sup>20</sup>*Modelling Manual Control of Straight Trajectories with a Perspective Flight-Path Display*. XVth European Annual Conference on Human Decision Making and Manual Control, 1996.
- <sup>21</sup>Warren, R. and Wertheim, A., editors, *Perception and Control of Self-Motion*, Lawrence Erlbaum Associates, Hillsdale, 1990.
- <sup>22</sup>Gibson, J. J., "Motion Picture Testing and Research," Tech. Rep. 7, Army Air Forces, Aviation Psychology Program Research Reports, 1947.
- <sup>23</sup>Gibson, J. J., *The Perception of the Visual World*, The Riverside Press, Cambridge, MA, 1950.
- <sup>24</sup>Gibson, J. J., "The Optical Expansion Pattern in Aerial Locomotion," *American Journal of Psychology*, Vol. 68, No. 3, 1955, pp. 448–484.
- <sup>25</sup>Gibson, J. J., Olum, P., and Rosenblatt, F., "Parallax and Perspective during Aircraft Landing," *American Journal of Psychology*, Vol. 68, No. 3, 1955, pp. 372–285.
- <sup>26</sup>Langewiesche, W., *Stick and Rudder*, chap. 5, McGraw-Hill Inc, 1944, pp. 276–277.
- <sup>27</sup>Wickens, C., "The Effects of Control Dynamics on Performance," *Handbook of Perception and Human Performance: Cognitive Processes and Performance*, edited by K. Boff, L. Kaufman, and J. Thomas, Vol. II, John Wiley and Sons, New York, 1986, pp. 39–1 to 39–60.
- <sup>28</sup>Sweet, B. T., *The Identification and Modeling of Visual Cue Usage in Manual Control Task Experiments*, Ph.D. thesis, Department of Aeronautics and Astronautics, Stanford University, Stanford, CA, 1999.

- <sup>29</sup>Sweet, B. T., “The Identification and Modeling of Visual Cue Usage in Manual Control Task Experiments,” Tech. Rep. TM-1999-208798, NASA, 1999.
- <sup>30</sup>Levison, W. H., “Measurement of Human Operator Response Behavior,” Tech. Rep. CR-166038, NASA, 1980.
- <sup>31</sup>Levison, W. H., “Some Computational Techniques for Estimating Human Operator Describing Functions,” Tech. Rep. NASA CP 2428, NASA, 1986.
- <sup>32</sup>Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., *Numerical Recipes in C*, chap. xx, Cambridge University Press, Cambridge, MA, 1992, pp. 657–661.
- <sup>33</sup>Mcruer, D. T. and Krendel, E., “Mathematical Models of Human Pilot Behavior,” Tech. Rep. AG-188, AGARD, 1974.
- <sup>34</sup>Greenwood, D. T., *Principles of Dynamics*, chap. 7, Prentice-Hall, Englewood Cliffs, 1988, pp. 354–358.
- <sup>35</sup>Schalkoff, R. J., *Digital Image Processing and Computer Vision*, chap. 2, John Wiley and Sons, New York, 1989, pp. 21–30.